

DOCUMENT RESUME

ED 199 083

SE 034 423

AUTHOR Warren, Jerry A.
TITLE Course Objectives and Planned Comparisons.
PUB DATE Aug 78
NOTE 15p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Behavioral Objectives; *College Mathematics; *Course Objectives; Course Organization; *Criteria; Curriculum Development; Educational Objectives; Higher Education; Mathematics Curriculum; Mathematics Instruction; Program Descriptions; *Program Development; *Specifications; *Statistics
IDENTIFIERS *Statistical Test Item Collection System

ABSTRACT

It is argued that course content should be specified as explicitly as possible in terms of what is expected of the student in the future. Four levels of details are addressed: (1) customary course objectives; (2) behavioral questions; (3) behavioral objectives; and (4) exams. The report described the Statistical Test Item Collection System (STICS), a support system consisting of a national file of statistical questions and answers and a series of computer programs. The three major points made by this document are: (1) statistics courses often vary wildly in content and requirements; (2) much of the variation can be traced to a failure to clearly specify what the course is to accomplish and what is to be required of students; (3) one way to provide the kind of specification needed is to record the behavioral objectives and the exam items that apply to a course. (MP)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

COURSE OBJECTIVES AND PLANNED COMPARISONS
Jerry A. Warren, University of New Hampshire
August, 1978

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Jerry A. Warren

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

ABSTRACT

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Once upon a time an investigator declared that he was ready to embark upon a marvelous experiment. He planned to use the most fashionable apparatus and techniques and anticipated a breathtaking advance in the frontiers of science. This investigator was saddened and disappointed by a friend who asked - What treatments will be tested? What comparisons will be made?

Once upon a time a statistician declared that he was ready to teach a marvelous course. He planned to use the most fashionable text, computer assisted head scratching, etc. and anticipated a colossal advance in the learning of statistics. The statistician was saddened and disappointed by a friend who asked - What changes will the course bring about in the future actions of students? What will be done to provide for consistency in what is asked of students if this course is repeated?

This paper is concerned with answers for the statistician's friend.

INTRODUCTION

A statistics course, like any other course, is an opportunity to affect the future behavior of students.

An experiment is an attempt to reveal some aspect of nature.

Even though these enterprises seem different, experimenters and the teachers of statistics courses have much in common. For example, either is likely to be "saddened and disappointed" when someone has the audacity to question the planning of one of their undertakings. Statistics courses often try to prepare experimenters for this kind of questioning by pointing out the need for planned comparisons. Yet, the teachers of statistics courses seldom are prepared for similar questions. The subculture of an experimenter is likely to include preparation for the question - "What are you trying to find out?". The subculture of teachers of statistics is not likely to include preparation for the question - "What do you want students to do when they leave your course?".

There are some important parallel needs common to conducting a course and conducting an experiment. Both should involve:

ED 199083

SE 034 423

- planning in considerable detail;
- recognition of the recursive nature of the activity; and
- recording of appropriate details.

Consideration of planned comparisons can be an antidote to such vague statements of experimental objectives as "Study the effects of supplemental organic matter on crop yields in the tropics". Addressing the question of comparing means (or equivalent questions for continuous treatment factors) should force us to check on correspondences among treatments to be tested, comparisons to be made, specific questions to be probed, and experimental technique. We make use of the wish to make comparisons as a way to clarify the focus of an experiment.

There should be a similar custom for statistics courses. Consideration of behavioral objectives can be an antidote to such vague statements of teaching objectives as "Study Regression Analysis". Eventually students leave a course. Before proceeding with the teaching of a course, we need to invest enough time to picture clearly what we want those students to do. Addressing the question of behavioral objectives should force us to check on the correspondence between those goals and the exams that students will take, a deed that requires identification of potential exam items.

The recursive nature of experimentation is widely recognized, e.g. Box (1). The customs of experimenters emphasize the need to record hypotheses, procedure, and results so that successors can build on earlier efforts. The recursive nature of teaching statistics courses ought to be equally widely recognized, but it seems to be generally ignored. The customs of teachers of statistics courses don't emphasize the need to record behavioral goals and exam items. They don't emphasize the need to record the details required for, say, the fourth iteration of a course to build on the preceding iterations.

Our customs are to prepare an experimenter for questions about the match between treatments and objectives. Our customs are to train experimenters so that there can be progress in a sequence of experiments. Our customs do not similarly provide for the teachers of statistics courses. The remainder of this paper is concerned with their needs.

SPECIFICATION OF A COURSE

Outlook

Before elaborating on the process of specifying (and respecifying) a statistics course, I will list a few of the viewpoints that affect my image of the process:

1. Degree of specification desirable. While I will describe a way of specifying the content of a course, I don't subscribe to the notion that any course should be fully specified. Depending on what is to follow a course and the importance of providing for the enthusiasms of instructors, I am inclined to think of specification targets that account for 20% to 80% of the course.

2. The allocation of effort by students. I view a typical student as one who does only what is necessary to get through a course. Usually what a student regards as necessary corresponds very closely to what is expected on exams. Exams dominate the allocation of student time, particularly when students have some reliable means of anticipating what will appear on exams.

3. Dependence on a Single Instructor. I tend to associate a high degree of specification for a course at least partially with the wish to avoid dependence on any single instructor. An introductory course that will be taught by many different people might have 70% of its content specified while an advanced course that is the specialty of one person might be 30% specified.

4. Course life span and review. I assume that effort spent on defining a course will be directed to courses with a life expectancy of 10-20 years. I also assume that there will be periodic reviews of heavily specified courses. I expect that the bulk of the effort spent in defining a course will occur during its first 2 or 3 years and that lesser amounts will be required at periods of 3-5 years.

5. Audience. Since I find that most of my contacts and clients come from 1-2 semester introductory level courses, I tend to think in terms of students who will confine their statistical training to those courses. I am not directly concerned with the training of future statisticians.

Objectives

In proposing specifications for a course I will be concerned with roughly four levels of details:

1. Customary course objectives,
2. Behavioral questions,
3. Behavioral objectives, and
4. Exam items.

Customary course objectives are the ritualistic statements that are used for course catalogs and administrative forms. Usually they are sufficiently vague so that they don't interfere with the specification of a course.

Behavioral questions are general questions concerned with what we want students to do after they have finished a course. When a student finishes Statistics XXX do we expect that the student will:

- perform long calculations and produce summaries of data?
- translate unfamiliar computing formulae into numerical results?
- use a specific computer program to perform calculations?
- verify the performance of any computer programs used to summarize results?
- derive theorems in new areas of application?
- recognize evidence that the assumptions for a standard technique have been violated?
- realize what information a client should supply to a consultant?
- etc.

Behavioral questions are the kinds of questions that we prefer not to raise because there is generally disagreement on which are important.

Behavioral objectives are statements of relatively specific kinds of behavior that we want students to display after finishing a course. Behavioral objectives are the answers for behavioral questions stated in terms of specific topics. Corresponding to a behavioral question about students trusting computer programs, I might propose a series of behavioral objectives such as -

Use one or more familiar data sets to check on the calculation of a regression program and the student's understanding of the labeling of output.

Use hand calculation of simple data sets to check on what is presented in computer output as a mean, standard deviation, standard error, etc.

The question of defining behavioral objectives is discussed much more fully in Mager (2).

Exam items are the questions asked of students. They usually are related to customary course objectives because those objectives ordinarily are sufficiently broad to cover almost any kind of exam. They are almost never explicitly related to behavioral questions or behavioral objectives because those matters are rarely addressed. Because behavioral objectives are rarely visible, it's not obvious that

they and their expression in exams change with instructors and with time. The course number and name stay the same, but the behavioral objectives and their reflections in exams often vary appreciably.

Suppose that a two-semester service course is to be taught for advanced undergraduates and graduate students in biology. I would suggest that the potential faculty for teaching that course and the faculty providing students for that course discuss behavioral questions. Following this the faculty who will teach the course would enumerate behavioral objectives and exam items and assign them to exam periods (usually 4-6 weeks) for the proposed course. Preferably, the number of exam items chosen would permit the use of sampling for preparing exams. Following suitable re-consultations and revisions, it should be possible to conduct the proposed course for 2-3 years during which exams (and the allocation of student effort) would be dominated by the mutually agreed upon collection of exam items. By the end of this time it will probably be obvious that adjustments are needed. Now the adjustments can be made on a course that has been specified in some detail. Adjustment now will tend to be concerned with re-examination of objectives and items and reallocation of items among exam periods.

The point of all this is that course content should be specified as explicitly as possible in terms of what we want students to do in the future. This requires recording behavioral objectives and exam items. Indicating that a course will follow a text or a syllabus ordinarily is not adequate. Behavioral targets are needed so that we can go through refinement cycles that adjust to our attempts at those targets.

THE STATISTICAL TEST ITEM COLLECTION SYSTEM, STICS

Present

STICS is a support system* that has been designed to assist in the specification of an introductory statistics course and the delivery of exams by sampling item collections. While STICS can provide help for these kinds of activity it cannot substitute for the most difficult part - the part that requires judgement in selecting behavioral objectives and recognizing when an exam item does or does not reflect an objective. The use of STICS to define a course has been described in some detail in Warren (4) and Michener (3).

The elements that currently make up STICS are still prototypes. Development has progressed to a point that makes it possible to see the general form of the computer programs and the exam item files, but not to the point where either is close to a finished state.

*This project is partially supported by National Science Foundation Grant SED 76-12191 A01.

STICS consists of a series of computer programs and a National File of statistical questions and answers. The National File now contains around 2000 items each of which includes an abstract, question, answer guideline, statistical key terms, and question characteristics. Statistical key terms indicate related topics such as regression, sampling, etc. Question characteristics cover such things as question type (essay, multiple choice, problem, ...), area of application, math prerequisite, etc. The National File now provides an uneven coverage of statistical topics. e.g. It has many entries concerned with the calculation of means and standard deviations and few entries for planned comparisons. The review pathway of an item in the National File is indicated in Appendix I.

The programs that make up STICS are self documented, conversational programs (a person using these programs should find all necessary instructions contained within the programs). The intent is that these programs should be useable by ordinary mortals - in distinction to computer addicts. There are 3 types of programs:

1. A program (SELECT) for extracting items from the National File - the purpose of this program is to allow a user to designate key term and question characteristic requirements to form a subset, a local file, that later might be used for preparing exams. SELECT allows a user to form relatively small question-answer files of 100-300 items. Files of this size are much easier to work with than the National File. A single course might make use of several local files corresponding to different portions of the course.

2. A set of Editing Programs - these programs allow a user to refine a subset that has been removed from the National File. They provide for adding, removing, or modifying items. In general, the editing programs allow a user to "improve" on what is in the National File. e.g. If a user finds that Tasmanian Devils occur in the items extracted from the National File and prefers questions that feature children, the editing programs provide a way to convert Tasmanian Devils into children (or vice versa).

3. An Exam Generating Program (BEGET) - this program provides for the use of random sampling to prepare exams and answer guidelines. If the exam items to be used for a 6 week period have been identified and put in a local file, then BEGET can be used to generate exams by sampling that file.

Future

During the next two years we expect that the major changes in STICS will involve enlarging and cleaning up the National File and broadening access to its programs. We particularly need help in improving the National File. Even though there were around 50 people who originally submitted items, the current collection has a high percentage of items from relatively few contributors. We would like to change this and ask your help in doing so. We would like to have, say, 5-20 items from a very large number of people. A sample of the format

that we would like followed appears in Appendix II. Among the ways that we suggest for choosing items are:

- Browse through the current National File and pick just one or two areas that are badly neglected. Then make up some questions that address those areas.

- Use problems or questions suggested by clients, students, newspaper articles, etc. to arrive at a small number of items to send in for the National File. e.g. Two questions suggested by recent consulting sessions are included in Appendix III.

We have no doubt that the National File needs improvement. We need your help to make those improvements. (We will be pleased to send microfiche copies of the National File to those who send useable items.)

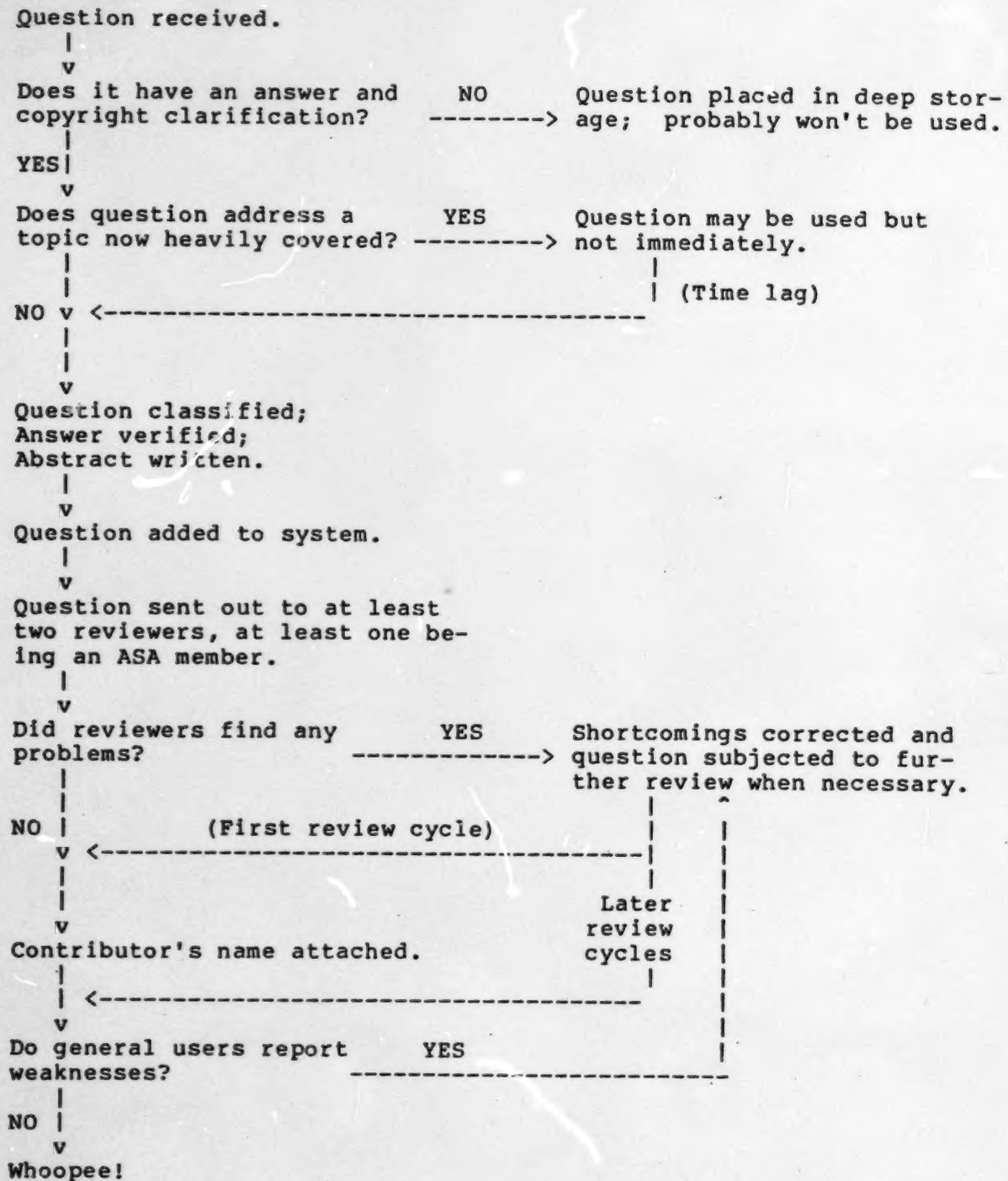
The STICS programs are written in FORTRAN and now run on DEC10 computers. They are scheduled for use on a DEC20 early this fall. During the coming year we plan to explore providing access to these programs via national networks such as EDUNET. In addition we plan to undertake installation on the major timesharing systems that run on IBM machines. While we plan to install the STICS programs on as many major systems as possible, we expect that those who are mainly interested in generating exams will often find it quicker to develop programs like BEGET that operate on local files.

Summary

- Statistics courses often vary wildly in content and requirements.
- Much of this variation can be traced to a failure to clearly specify what the course is to accomplish and what is to be required of students.
- One way to provide the kind of specification needed is to record the behavioral objectives and the exam items that apply to a course.
- STICS is a collection of computer programs and question-answer files that can be used to support this kind of specification.

APPENDIX I

Review of Items in the System:



APPENDIX II

Format for Items for the National File

A copy of the form suggested for submitting items to the National File follows. We have separated essential elements from helpful elements to make it easier for those who wish to submit items without worrying about abstracts and key terms. We can supply the latter, but would prefer that the original contributor make an attempt at doing so. (It's not necessary to refer to our key term list. List the terms that you would like to have associated with an item.)

A copy of our Copyright Clarification Form also follows. Since copyright clearance is needed if items are to appear in the National File, we do not propose to reproduce or distribute any items for which there are conflicting copyright claims. Other grounds for not using items are: lack of clarity, incorrect statistical content, or occurrence in a category that already has many items.

Contributors of items for the National File who wish to have notice of acceptance sent to a local administrator should include the name and address of that administrator with their name.

QUESTION SUBMITTAL FORM

I. Essentials (If space is inadequate, attach additional pages.)

A. Contributor:

Name _____ Date _____

Address _____

B. Question:

C. Answer:

II. Helpful

A. Abstract (1-2 sentences)

B. Suggested Key Terms (Guess at 5 words that someone else might think of looking up when searching for this item.)

C. Name and Address of administrator if you wish notice of acceptance sent.

UNIVERSITY OF NEW HAMPSHIRE STATISTICS TEACHING PROJECT*
DIAGNOSTIC SYSTEM

Clarification of Copyright Status

1. Material acquired: _____

2. Material received from NAME: _____
UNIVERSITY: _____
ADDRESS: _____

3. Is copyright of this material owned by someone other than yourself?
No _____ Yes _____ If yes, please identify the source of the materials
and the copyright owner (if any):
NAME: _____
ADDRESS: _____

4. Will you give permission to publish in our catalog, reprint, and
distribute the material listed above?

_____ No

_____ Yes, the materials listed above may be published in your
catalog under the proviso that materials not currently
subject to copyrights will be copyrighted on my behalf,
and they may be reprinted and distributed by the Univer-
sity of New Hampshire Statistics Teaching Project mem-
bers, with a general permission for other individuals to
reproduce them.

Signature

Date

Please return this form to: Alvin Bugbee, Manager
Diagnostic System
304 McConnell Hall
University of New Hampshire
Durham, New Hampshire 03824

*This project is partially supported by National Science
Foundation Grant SED 76-12191 A01.

APPENDIX III

Two Questions Suggested by Consulting Sessions

1. Q. You have just completed summarizing some data on a computer and have looked at the analysis of variance table included in the printout. One of the entries in the sum of squares column is negative. What does that mean?
 - a. The source of variation associated with that sum of squares had a negative effect on response.
 - b. The data has not been properly scaled and has to be normalized.
 - c. A transformation is needed.
 - d. Statistics is a hopeless subject.
 - e. There is something wrong with the computer program.

A. e. Since by definition a sum of squares must be either zero or positive, the computer program almost certainly contains a blunder when a sum of squares values is negative. Whoever is responsible for the program that was used should be notified that something is wrong.
2. Q. A computer programmer was concerned about the length of time that would be required to print exam questions and answers at a terminal (a printing device like a typewriter). The programmer knew:
 - that the terminal was capable of printing 30 characters per second;
 - the number of lines (maximum length: 80 characters) of text that made up each question and answer;
 - the identity of the questions to be printed at any particular time; and
 - that question selection for a single exam was a random process.

Accordingly, the programmer estimated the time required to print a set of questions and answers by multiplying the total number of lines to be printed by 80 and then dividing the result by 30 to arrive at an estimated print time in seconds. Later the programmer was informed that his estimates for print time were always too high. What do you recommend that the programmer do to improve his estimation scheme?

- A. A variety of suggestions should be offered in response to this question. Some should focus on the distribution of characters per line and the use of something other than the maximum length as a basis for estimation. (This assumes that the printing device recognizes short lines and does not always attempt to

print 80 characters.) My preference would be to form a series of exams covering the anticipated range of usage and using random selection as much as possible. I would print these exams on the terminal and hope that a regression involving:

Y: Print time; and
X: Total number of lines

would provide an adequate approximation for estimating print time. If not, I would then be inclined to examine use of such variables as question length, answer length, computer load, etc.

REFERENCES

1. Box, G.E.P. (1976) "Science and Statistics", Journal of the American Statistical Association, 71:791-799.
2. Mager, R.F. (1962) Preparing Instructional Objectives, Belmont: Fearon Publishers.
3. Michener, R.D., Bugbee, A.D., and Warren, J.A. (1978) "Automating Exams for a Statistics Course - II. A Case Study", Association for Educational Data Systems Proceedings, Atlanta, Georgia.
4. Warren, J.A., Michener, R.D., and Bugbee, A.D. (1978) "Automating Exams for a Statistics Course - I. The Beginning of a Case Study", Association for Educational Data Systems Proceedings, Atlanta, Georgia.